

온톨로지 기반 영상이해 시스템

Ontology-based Image Understanding Systems

이인근 서석태 정혜천 손세호* 권순학**

In K. Lee, Suk T. Seo, Hye C. Jeong, Seo H. Son*, Soon H. Kwon**

영남대학교 전기공학과, *담 엔지니어링

요 약

온톨로지는 공유된 개념과 그 개념들 사이의 관계로 표현된다. 이러한 온톨로지를 사용하여 인간과 시스템에 대한 지식의 공유에 관한 연구가 활발히 이루어져 왔다. 예를 들면, 온톨로지의 설계 및 구축에 의한 영상이해를 들 수 있다. 그러나 온톨로지에 기반한 영상이해 방식 중 대부분의 기존 방식은 개념적인 연구에 그칠 뿐 구체적인 방법을 제시하지는 못하였다. 본 논문에서는 온톨로지로 표현된 지식에 근거하여 영상을 이해하는 다음과 같은 영상이해 프로세스 및 시스템을 제안한다. i) 특정 분야의 지식을 온톨로지로 표현하고, ii) 영상 처리 및 분석 과정을 통해 영상을 구성하는 객체들의 특징을 추출하며, iii) 객체의 특징으로부터 객체의 개념을 해석하고, iv) 온톨로지 추론을 통해 영상 해석 과정에서의 애매성을 줄인다. 제안된 영상 이해 프로세스에 기반하여 영상이해 시스템을 구축하고, 특정 분야에서의 실험을 통하여 제안된 프로세스와 시스템의 효용성을 확인한다.

키워드 : 영상이해, 온톨로지, 영상 해석

Abstract

Ontology is represented by the shared concepts and relations among those. Many studies have been actively working on sharing human's knowledge with that of systems by using it. For a typical example, there is the design and implementation of ontology system for image understanding. However conventional studies on ontology-based image understanding have proposed not concrete methods but conceptual idea. In this paper, we propose an ontology-based image understanding system with following four processes: i) knowledge representation of a specific domain by the ontology, ii) feature extraction of objects through image processing and image analysis, iii) image interpretation by object features, and iv) reduction of ambiguity existing in image interpretation by ontology reasoning. We implement an image understanding system based on the proposed processes, and show the effectiveness of the proposed system from experimental results in a specific domain.

Key Words : image understanding, ontology, image interpretation

1. 서 론

컴퓨터가 발달함에 따라서 영상에 관한 연구가 점차 활발히 행해지고 있으며, 이진화, 에지 추출, 노이즈 제거 등의 단순 영상처리를 넘어서 지문인식, 얼굴인식, 등과 같은 생체 인식에 관한 연구가 많이 이루어지고 있다. 이러한 연구들의 근본적 목표는 주어진 영상 내에서 객체를 인식하고 객체 사이의 관계를 추출하여 영상을 이해하는 것이다.

이러한 목적에 의해서 SIGMA, SCHEMA, SYGAL 그리고 PROCAL과 같은 다양한 지식을 이용한 영상이해 시스템들이 제안되었다[1-3]. 또한 최근 들어서 Bayesian network를 기반으로 한 영상이해 시스템이 제안되었다[4]. 그러나 이러한 영상이해 시스템은 실제적인 구현에 있어서 한계가 있다. 시스템 구현의 한계성을 극복하기 위해서 지식에 기반한 영상이해에 대한 연구가 이루어지고 있다. Maillot은 색 개념

(color concept), 질감 개념(texture concept) 그리고 공간적 시각 개념(spatial-temporal concept)을 사용하여 정의역에 독립적인(domain independent) 지식 베이스를 구축한 영상이해 시스템을 제안하였다[5,6]. 또한 Son 등은 온톨로지의 의미학적 분류와 구성학적 분류를 이용하여 특정 영역의 개념들에 대한 정의역에 독립적인 지식베이스를 구성하는 영상이해 시스템을 제안하였다[1].

온톨로지(ontology)란 특정 영역의 개념(concept)을 정의하고 개념의 속성(property), 개념들의 관계(relation)에 대한 지식 표현으로 정의된다[7]. 현재 지식 구성이 가능한 온톨로지를 이용한 많은 응용 시스템이 개발되고 연구되고 있다. 그리고 본 논문에서는 특정 분야의 온톨로지를 OWL (Ontology Web Language)로 구성하고, 영상으로부터 객체의 속성을 추출하여 영상을 해석하고 이해하는 영상이해 프로세스와 시스템을 제안한다.

본 논문에서 정의하는 영상이해(image understanding)는 영상에 포함되어 있는 객체(objects)의 부분 이미지로부터 해당 객체를 인식하고, 인식한 객체의 행동이나 객체와 객체 사이의 관계를 규정하여 전체 영상의 의미를 해석하고 묘사(description)하는 것으로 규정한다. 즉, 영상이해의 범위를

**Corresponding author

접수일자 : 2007년 1월 5일

완료일자 : 2007년 5월 25일

i) "객체의 개념 인식", ii) "객체 사이의 관계 규정", iii) "영상의 의미 해석과 묘사"로 구분하여 정의한다. "객체의 개념 인식"은 영상내의 관심 객체의 영역을 분리하고, 분리된 객체 영역에서 추출할 수 있는 특징(feature)을 기준으로 객체의 개념을 해석한다. 이 과정에서, 객체 영역에서 추출한 특징을 미리 학습된 영상 객체의 특징과 비교하고, 온톨로지에 기반한 추론에 의해 객체의 개념 해석 과정에서의 애매성을 줄인다. "객체 사이의 관계 규정"은 영상에서의 객체 관계를 표현한 온톨로지를 기반으로 하여 두 객체의 상호 관계와 객체의 동작을 해석하는 것이다. 그리고 "영상의 의미 해석과 묘사"는 영상의 객체 개념과 객체간의 관계를 통해 전체 영상이 의미하는 바를 해석하여 그 결과를 정형적으로 묘사하는 것이다.

본 논문에서는 영상이해 범위 중 "객체의 개념 인식"을 연구 범위로 정한다.

2. 영상이해 프로세스

본 논문에서 제안하는 영상이해 과정은 다음과 같다. i) 특정 분야의 지식을 온톨로지 표현하고, ii) 영상 처리 및 분석 과정을 통해 영상을 구성하는 객체들의 속성을 추출하며, iii) 온톨로지 추론을 통해 객체의 속성으로부터 객체를 정의하여 영상을 해석한다.

영상이해 프로세스는 그림 1과 같다.

[영상이해 프로세스]

[step1] 영상에서 객체를 분리한다.

[step1-1] 객체가 없으면 [step8] 실행한다.

[step2] 객체에서 특징 값을 추출한다.

[step3] 객체의 특징 값으로 객체를 해석한다.

[step3-1] 추출한 특징 값과 학습영상 객체 특징 값과의 유사도를 측정한다.

[step3-2] 객체 개념들의 유사도가 모두 임계값 이하일 경우 [step1] 실행한다.

[step3-3] 객체 개념들의 유사도가 임계값 이상인 개념 후보를 선택한다.

[step4] 객체 해석 결과가 내부 객체의 해석 결과인지 판단한다.

[step4-1] 내부 객체의 해석 결과인 경우 객체 정보 임시 저장소에 저장되어 있는 객체 개념 후보들을 참조하여 객체를 재해석 하고, [step4] 실행한다.

[step5] 객체 정보 임시 저장소에 저장된 객체 개념 후보의 개수가 1개인 경우 [step7] 실행한다.

[step6] 객체 영역 내의 내부 객체를 분리한다.

[step6-1] 내부 객체가 있으면 [step2] 실행한다.

[step6-2] 내부 객체가 없으면 개념 후보에서 최상위 개념 후보를 선택한다.

[step7] 해석한 객체의 개념을 객체 개념 임시저장소에 저장하고 [step1] 실행한다.

[step8] 객체들의 개념을 조합하여 전체 개념을 추론한다.

[step9] 영상 해석을 마친다.

[step1]에서는 입력 영상에서 분리 가능한 객체를 분리하여 배경을 제거한 새로운 객체 영상을 생성한다. [step2]에서는 객체 영상으로부터 색상, 형태, 질감 등의 특징 값을 분석하고, [step3]에서는 온톨로지에 저장되어 있는 학습된 영상

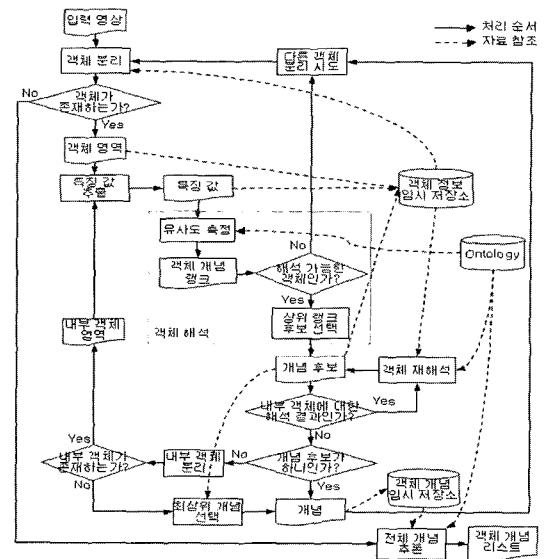


그림 1. 영상이해 프로세스

Figure 1. Image understanding process

개체들의 특징 값과 비교하여 객체의 개념을 해석하고 객체 개념의 후보들을 생성한다. 이때, 유사도가 임계값 이상인 개념 후보를 선택하고, 전체 후보들의 유사도가 임계값을 넘지 못하는 경우에는 객체를 해석할 수 없다고 판단한다. [step4]에서는 객체 해석 결과가 내부 객체에 대한 해석 결과일 경우 추론을 통해 객체의 개념을 재해석 한다. 그리고 객체의 개념 재해석 결과로 얻은 개념 후보가 2개 이상인 경우에는 객체 해석에 대한 애매성을 줄이기 위해 [step6]에서 내부 객체의 해석을 시도한다. 입력 영상으로부터 더 이상 분리 가능한 객체가 존재하지 않을 경우, [step8]에서 분리한 객체들의 개념을 조합하여 개념을 추론한다. 영상이해 과정에서 분리한 객체들의 정보는 '객체 정보 임시 저장소'나 '객체 개념 임시 저장소'에 저장하여 영상이해에 활용한다.

3. 영상이해 시스템

3.1. 영상이해 시스템의 구성

제안한 영상이해 프로세스에 따른 영상이해 시스템은 전문가에 의해 구축된 "지식 베이스", 입력 영상에서 해석하고자 하는 객체의 "영역 선택", 선택된 영역에서 객체의 특징을 추출하는 "객체 특징 추출", 추출된 객체의 특징과 지식 베이스를 이용하여 객체의 개념을 추론하는 "객체 해석"으로 구성되어 있다. 그림 2는 제안한 영상이해 시스템의 구조를 보인다.

3.2. 영상이해 시스템 지식베이스

3.2.1. 온톨로지 구성

영상이해의 지식베이스는 온톨로지를 기반으로 하며 OWL로 구성한다. OWL은 정보의 표현 뿐 아니라, 정보의 내용을 어플리케이션에서 직접 처리할 수 있도록 설계된 언어이다[8]. OWL은 class, property, individual, restriction으로 구성할 수 있다. 각 개체는 기본적인 속성 외에도 영상해석을 위해 정의한 특징 속성을 부여한다. 그림 3과 같이 온톨로지는 비슷한 개체들이 모여 하나의 개념을 형성하며, 개

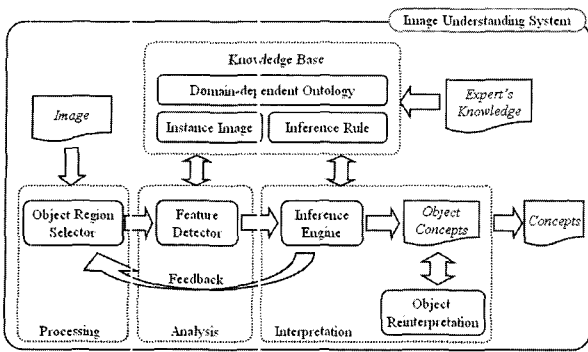


그림 2. 영상이해 시스템 구조

Figure 2. The structure of the image understanding system

념은 그 의미에 따라 계층으로 구성된다. 그리고 객체의 속성에 따라 객체와 개념간의 관계가 형성된다.

본 논문에서는 그림 4와같이 실제 영상(real image)을 객체로 하고, 영상개체(image instance) 속성으로는 영상개체 내부의 객체에 대한 색상(color), 형태(shape), 질감(texture)과 같은 시각적 속성을 사용한다.

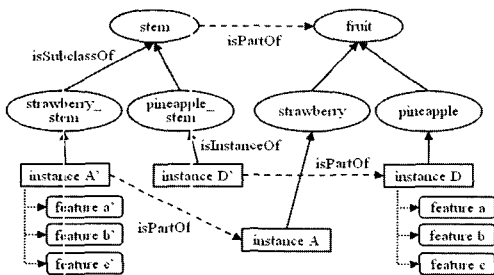


그림 3. 온톨로지 구성 예

Figure 3. An example of Ontology

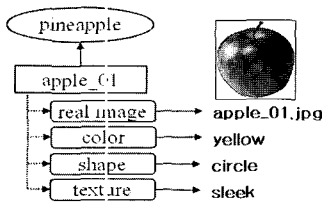


그림 4. 개체의 속성 예

Figure 4. An example of instance properties

3.2.2. 온톨로지 영상 학습

본 논문에서의 영상 학습은 사람이 사물을 시각을 통해 인지하고 학습하는 방법을 모델로 하였다. 사람은 과거에 보았던 사물들의 특징들에 대한 기억과 처음 본 사물의 특징을 비교하여 그 사물의 개념을 판단한다. 이와 같이 동일한 개념을 표현하는 영상개체라 하더라도 입력 영상개체의 특징이 기존에 학습한 영상개체의 특징과 전체적으로 일치하지 않을 수도 있다. 따라서 입력 영상개체의 속성이 학습 영상개체의 속성 범위 내에 속하는 경우 동일 개념의 영상으로 판단한다. 예를 들어, 표 1과 같이 동일한 “사과”라는 개념이지만 시각적 특징이 다른 두 영상개체 A와 B를 “사과”라는 개념과 연결 지어 인지하고 학습한다. 만일 새로운 영상개체 C의

특징이 기존에 학습한 영상개체 A와 B의 특징과 전체적으로 일치하지 않고 그들의 특징 범위 내에 속할 경우, 영상개체 C도 “사과”의 개념 범위에 속할 수 있다고 판단한다.

본 논문에서는 개념별로 몇 개의 영상개체를 분석하여 객체 해석을 위한 개념의 속성 범위를 정하는 것을 “영상 학습”이라고 정의한다.

표 1. 영상의 학습과 해석 예

Table 1. An example of image learning and interpretation

[개념 학습]
개체A=“빨강고 원형의 사과”
개체B=“파랗고 타원형의 사과”
[개념 속성 정의]
사과= \forall 형태.(원 \cup 타원) \cap \forall 색.(파랑 \cup 빨강)
[개체 입력]
개체C=“파랗고 원형의 사과”
[개체 추론]
개체C.형태 \in 사과.형태 \cap 개체C.색 \in 사과.색
[개체 해석]
개체C \in “사과”

3.3 영상 처리

온톨로지 지식베이스가 구성되면 실제 이미지를 입력하여 이미지로부터 관심 객체를 추출하고 추출된 객체의 속성을 분석한다.

영상에서 해석하고자 하는 객체의 특징을 추출하기 위해서는 원 영상에서 관심 객체만을 분리함으로써 특징 추출 과정에서 관심 객체 이외의 정보를 제거하는 것을 목적으로 한다. 즉, 원 영상으로부터 영상 처리 과정을 통해 관심 객체만을 분리하여 영상 해석의 애매성을 줄인다.

3.3.1. 객체 분리

영상에서 관심 있는 하나의 객체를 중심으로 그 이외의 부분들은 배경으로 판단하는 것을 객체 분리의 기본 개념으로 정의한다. 이 개념에 의하면 영상에서 단·복수의 객체와 배경을 분리해 내고, 객체 내부에서 다시 관심 객체와 그 이외의 객체를 분리해 내는 과정을 반복하여 다수의 객체를 분리해 낼 수 있다. 그러나 객체 분리 개념으로 실제 영상에서 객체를 분리할 경우 해석해야할 객체의 수가 폭발적으로 늘어날 가능성이 있다. 따라서 분리할 수 있는 가장 큰 객체로부터 내부의 객체를 점차적으로 분리하여 객체를 해석한다.

영상에서 표현되는 객체의 형태는 다음과 같다.

- (a) 1개의 개념이 1개의 객체로 분리될 수 있는 경우
- (b) 1개의 개념이 다수의 객체로 분리될 수 있는 경우
- (c) 다수의 개념이 1개의 객체로 분리될 수 있는 경우
- (d) 다수의 개념이 다수의 객체로 분리될 수 있는 경우

(a)는 1개의 개념이 1개의 객체 영역으로 분리될 수 있고, 분리한 객체 내부에서 여러 개의 또 다른 객체를 분리할 수 있는 경우를 말한다. 예를 들어, 딸기의 객체 내부에서 딸기 씨앗 객체를 분리할 수 있는 경우이다. (b)는 분리할 수 있는 다수의 객체가 모여 하나의 개념을 나타내는 경우이다. 예를 들어, 파인애플의 꼭지 객체와 파인애플 몸통 객체를 각각 분리할 수 있지만 두 객체가 모여 하나의 개념을 나타내는 경우이다. (c)는 노란색의 꿀 객체가 쌓여 있는 것과 같이 속

성이 비슷한 객체들이 모여 있어 객체 사이의 경계가 불분명하여 전체가 하나의 객체로 분리되는 경우이다. (d)는 과일 바구니의 다양한 과일들처럼 속성이 다른 다양한 객체가 모여 있는 경우이다. 본 논문에서는 (a)와 (b)의 경우만을 고려한다. 그림 5는 원 영상으로부터 관심 객체 영역을 추출하고, 추출한 영역 내에서 새로운 객체를 분리하는 과정을 보인다.

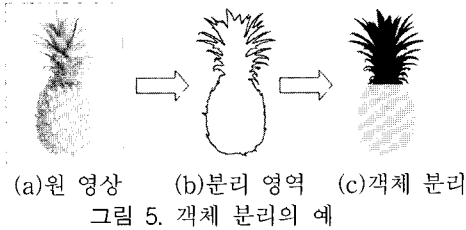


Figure 5. An example of object separation

3.4 영상 분석

영상 분석은 영상 처리 과정에서 관심 객체를 분리하고, 해석하기 위한 판단의 기준이 되는 특징을 추출하는 과정이다. 영상 객체에서 정의할 수 있는 특징은 내용기반 영상검색에서 많이 사용하고 있는 색상, 형태, 질감을 이용하며, 그 외의 특징도 필요에 따라서 추가 정의할 수 있다.

3.4.1. 색상 분석

분리된 객체로부터 색상을 분석한다. 식 (1)과 같이 색상의 범위를 RGB 각 채널별로 4개의 범위를 지정하고 총 64가지의 색상 그룹(Color Group)을 생성한다.

광원의 영향을 최소화하기 위해 RGB색상 채널의 각 채널 값인 R, G, B 를 식 (2)와같이 정규화 한다. 정규화한 N_R, N_G, N_B 에 따라서 식(3)의 조건에 따라 픽셀을 포함할 색상 그룹을 선정한다. $G_{rgb} = \{G_{111}, \dots, G_{444}\}$ 으로 표현되는 64개의 색상 그룹을 기준으로 객체 색상의 히스토그램을 구한다.

$$Color\ Group = \{G_{rgb} | 1 \leq r, g, b \leq 4, r, g, b: \text{자연수}\} \quad (1)$$

$$N_R = \frac{R}{R+G+B} \quad N_G = \frac{G}{R+G+B} \quad N_B = \frac{B}{R+G+B} \quad (2)$$

$$\begin{cases} \text{if } 0 \leq N_R < 0.25 & \text{then } r = 1 \\ \text{if } 0.25 \leq N_R < 0.5 & \text{then } r = 2 \\ \text{if } 0.5 \leq N_R < 0.75 & \text{then } r = 3 \\ \text{if } 0.75 \leq N_R \leq 1 & \text{then } r = 4 \end{cases} \quad (3)$$

(N_G, N_B 에 대해서 식 (3)과 같은 방법으로 g, b 를 구한다.)

그리고, 색상은 다음과 같이 표현될 수 있다. 예를 들어, 그림 6과 같이 회색(gray)을 표현하는 방법으로 검정색(black)과 백색(white)의 픽셀을 혼합하여 전체 색상을 표현하거나 전체 픽셀을 회색으로 표현하기도 한다. 이 경우 같은 회색일지라도 픽셀의 색상 분포는 다르다. 따라서 식 (4)와같이 일정 크기의 윈도우(window) 내의 색상 값을 평균한 색상을 윈도우 중심 픽셀의 색상으로 사용하여 색상 표현 방법에 대한 차이를 최소화한다.

$$n_z = \frac{\sum_{i=1}^N \sum_{j=1}^M Z_{ij}}{n \times m} \quad (4)$$

n_z : $m \times n$ 윈도우 내의 평균 색상 값
 N : $m \times n$ 윈도우 내의 전체 픽셀 수
 Z_{ij} : $m \times n$ 윈도우 내의 (i, j) 번째 픽셀의 색상 값

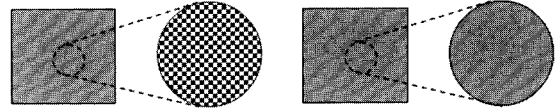


Figure 6. Color expression of gray

3.4.2. 형태 분석

객체의 형태 특징을 분석하기 위해서 영상에서 분리된 객체로부터 형태 정보를 추출한다. 본 논문에서는 객체의 형태를 표현하기 위해서 식 (5)와같이 객체의 중심(x_c, y_c)에서 외곽(x_s, y_s)까지의 거리를 일정 각도(d')씩 회전하면서 측정한다. 그리고 식 (6)과같이 정규화된 형태 분석 값 S_d 를 구한다.

$$s_d = \sqrt{(x_c - x_s)^2 + (y_c - y_s)^2} \quad (5)$$

$$S_d = \frac{s_d}{\arg \max_{0 < d' \leq 360} (s_d)} \quad (6)$$

예를 들어, 그림 6(a)와같이 원형의 경우는 거리가 일정하여 일자 모양의 형태 특징을 얻을 수 있다. 그리고 그림 6(b)와같이 사각형의 경우는 각도에 따라 물결 모양의 형태 특징을 얻을 수 있다.

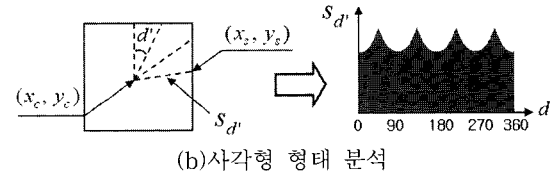
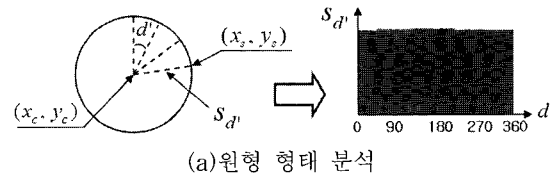


Figure 7. An example of shape analysis

3.4.3. 질감 분석

질감을 분석하여 특징을 얻기 위해서는 그림 8과같이 객체의 Hue 채널과 Gray 채널에서 종축으로 일정 간격의 기준(l)을 선정하고, 각 기준별로 횡축 P 개 픽셀 열에서 Hue 채널 값(H_p)과 Gray 채널 값(G_p)을 분석하여 식 (7)과같이 각 기준에서의 질감 분석 값(T_l)을 구한다.

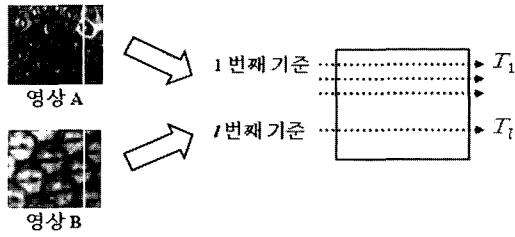


그림 8. 질감 분석 예

Figure 8. An example of texture analysis

$$H_{average} = \frac{\sum_{p=1}^P H_p}{P}, \quad G_{average} = \frac{\sum_{p=1}^P G_p}{P}$$

$$H_{distribution} = \frac{\sum_{p=1}^P (H_p - H_{average})^2}{P}$$

$$G_{distribution} = \frac{\sum_{p=1}^P (G_p - G_{average})^2}{P}$$

$$T_l = (2 \times H_{distribution}^l) \times G_{distribution}^l \quad (7)$$

그림 9는 그림 8에서 하나의 기준(l)에 대해 횡축 픽셀열에서의 Hue 채널에 대한 질감 특징을 나타낸다. 예를 들어, A영상처럼 표면이 매끄럽고 색상의 변화가 단조로운 경우는 그림 9(a)와같이 특징 값이 균일하며, B영상처럼 표면이 불규칙한 경우는 그림 9(b)와같이 특징 값이 불규칙하다.

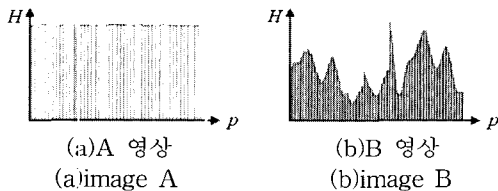


그림 9. Hue 채널에 대한 영상의 질감
Figure 9. Image texture by Hue channel

3.5. 영상 해석

영상 해석은 객체 영역으로부터 추출한 특징 값을 학습 영상의 객체 특징 값과 비교하여 객체의 개념후보를 제시하는 과정이다.

본 논문에서는 객체의 특징 속성을 색상, 형태, 질감으로 정의하였다. 따라서 3가지 객체의 특징 값을 기준으로 학습 객체의 특징 값과의 유사도(similarity)를 측정하고, 각 특징 속성에서의 유사도를 조합하여 객체의 개념을 판단한다.

3.5.1. 색상 유사도 측정

색상간의 유사도를 측정하기 위한 유사도 측도(similarity measure)로 색상 그룹별 차이를 도입하여 식 (8)과같이 정의한다.

$$S_{Color} = \frac{1}{\sum_{r=1}^4 \sum_{l=1}^4 \sum_{b=1}^4 |G_{rgb}^i - G_{rgb}^t| + 1} \quad (8)$$

G_{rgb}^t : 테스트 영상 색상 그룹

G_{rgb}^i : 학습한 영상객체의 색상 그룹

3.5.2. 형태 유사도 측정

형태 유사도 측정할 때 같은 형태의 객체라 하더라도 특징 값 분석 기준이 다를 경우 특징 값은 완전히 달라질 수 있다. 따라서 그림 10과같이 비교하려는 두 객체의 각도 차이에 의한 문제를 해결하기 위해 해석하려는 객체를 회전하여 얻은 형태 특징 값을 학습 객체의 형태 특징 값과 비교하여 유사도를 구한다. 그리고 식 (9)와같이 객체를 회전하여 얻은 유사도 중 가장 큰 유사도를 두 객체의 형태 유사도로 선택한다. R 은 테스트 영상 객체 r' 각도로 회전한 횟수이고, D 는 영상 객체의 d' 각도에 대한 특징 값의 개수이다.

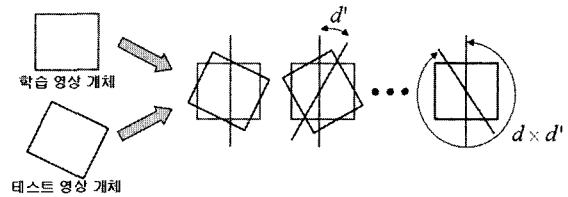


그림 10. 형태 유사도 측정 예

Figure 10. An example for similarity measurement of shape

$$S_{Shape} = \arg \max_{1 \leq r \leq R} \left(\frac{1}{\sum_{d=1}^D (S_d^i - S_{d,r}^t)^2 + 1} \right) \quad (9)$$

$$R = \frac{360}{r'} \quad \left(R \text{ and } r' : \text{자연수} \right)$$

$$0 < r' \leq 360$$

$$D = \frac{360}{d'} \quad \left(D \text{ and } d' : \text{자연수} \right)$$

$$0 < d' \leq 360$$

$S_{d,r}^t$: r 번째 영상객체 회전에 대한 테스트 영상의 d 각도에 서의 형태 특징 값

S_d^i : 학습 영상객체의 d 각도의 형태 특징 값

3.5.3. 질감 유사도 측정

객체의 각도에 따라서 질감의 유사도 또한 변하게 된다. 그리고 객체의 위치에 따라 다른 질감을 나타낼 수 있다. 그러므로 형태 유사도 측정 방법과 같은 방법으로 질감 유사도를 측정한다. 질감 유사도 측정은 식 (10)과 같다.

$$S_{Texture} = \arg \max_{1 \leq r \leq R} \left(\frac{1}{\sum_{d=1}^D (T_l^i - T_{l,r}^t)^2 + 1} \right) \quad (10)$$

T_l^i : 학습 영상객체의 l 기준에서의 질감 특징 값

$T_{l,r}^t$: r 번째 영상객체 회전에 대한 테스트 영상의 l 기준에 서의 질감 특징 값

3.5.4. 유사도 조합

객체의 각 속성이 영상에서 표현될 때는 주변 환경에 따

라 그 특징 값이 달라진다. 색상은 광원의 위치, 조명의 색, 그림자에 따라서 그 특징 값이 달라질 수 있다. 모양은 피사체의 각도에 따라서 그 모양이 달라지고, 다른 객체에 가려지는 경우 그 모양을 판단할 수 없다. 질감은 광원, 피사체의 각도 등에 따라서 질감이 달라질 수 있고, 같은 피사체라 하더라도 부위에 따라서 다른 질감을 나타낼 수 있다. 따라서 영상 해석의 목적에 따라 각 속성의 유사도를 적절히 조절하여 객체 간 유사도를 구한다. 식 (11)과같이 각 특징 유사도 측정값에 가중치를 부여하여 최종 유사도(S_T)를 구한다.

$$S_T = \alpha \times S_{Color} + \beta \times S_{Shape} + \gamma \times S_{Texture} \quad (11)$$

$$(\alpha \geq 0, \beta \geq 0, \gamma \geq 0)$$

3.6. 영상 객체 재해석

영상 객체에 대한 재해석은 i)영상 해석 과정에서의 애매성을 줄이기 위한 방법과 ii)해석한 객체들의 조합으로 하나의 개념을 나타내기 위한 방법으로 정의한다.

영상 객체 해석 과정에서 애매성이 너무 클 경우 객체 내에 존재하는 내부 객체를 분리하여 다시 영상을 해석한다. 즉, 그림 11과같이 객체 개념과 내부 객체 개념과의 관계를 온톨로지를 통해 추론하여 객체의 개념을 정의한다.

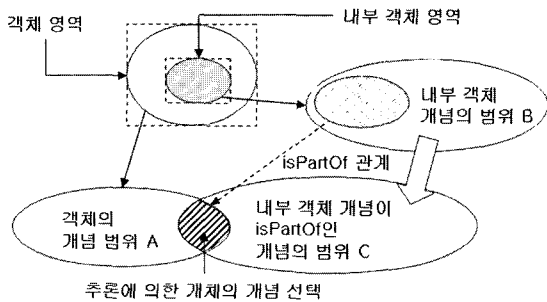


그림 11. 영상 객체 재해석 개념도
Figure 11. A configuration of image object reinterpretation

전체 객체 영역에 대한 영상 해석 결과인 객체의 개념 범위 A에서 개념 범위가 넓어 객체의 정확한 개념을 판단하기 어려운 경우, 객체 내부에 존재하는 독립 객체에 대한 영상 해석을 한다. 내부 객체에 대한 영상 해석 결과인 내부 객체의 개념 범위 B를 얻는다. 내부 객체 개념의 범위 B 또한 개념 범위가 넓어 내부 객체의 정확한 개념을 판단하기 어려울 수 있다. 이 경우 객체와 내부 객체의 개념 관계는 "isPartOf"이다. 즉, 내부 객체가 객체의 한 부분으로 존재한다는 온톨로지 지식 정보를 참조하여 내부 객체의 개념들과 "isPartOf"의 관계에 있는 객체의 개념 범위 C를 정의하고, 개념 범위 C와 개념 범위 A의 공통 개념 범위가 객체의 개념일 가능성이 높다는 가정이다. 이런 방법으로 영상 해석 과정에서의 애매성(ambiguity)을 줄인다.

그리고, 영상에서 나타나는 객체 자체가 독립적으로 의미를 지닐 수 있다. 그러나 객체들이 모여 더 큰 객체를 나타낼 경우 그 상위 개념(upper concept) 또한 영상에서 중요한 객체일 수 있다. 즉, 그림 12와같이 "사과", "바나나", "딸기" 등의 각 객체의 개념들이 모여 "과일바구니"와 같은 공통된 개념을 나타낼 수 있다. 따라서 추출한 객체의 개념을 조합하여 넓은 범위의 개념을 추론한다.

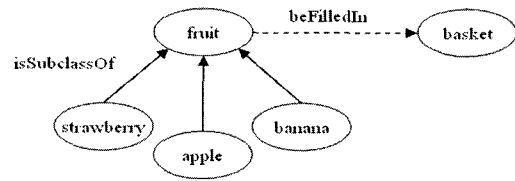


그림 12. 객체 개념 조합 예
Figure 12. An example of object concept composition

4. 모의실험

제안된 영상이해 프로세스와 시스템의 효용성을 검증하기 위하여, 영상이해 프로세스에 기반한 영상이해 시스템을 구축하고, 이를 이용하여 영상이해 실험을 수행하였다.

Protégé를 이용한 OWL 형태의 식물 분야(plant domain) 온톨로지를 구축하였다. 온톨로지 구축 시 WordNet[9]을 기반으로 식물 분야의 상위 개념의 계층 구조를 형성하였다. 그리고 학습에 이용한 영상을 분석하여 각 특징 속성의 특징 값을 추출하고, 학습 영상과 학습 영상개체에 포함된 객체에서 추출한 특징 값을 각각 개념의 영상개체와 영상개체의 특징 속성 값으로 설정하였다. 그림 13은 Protégé[10]를 이용하여 구현한 OWL 온톨로지를 보인다.

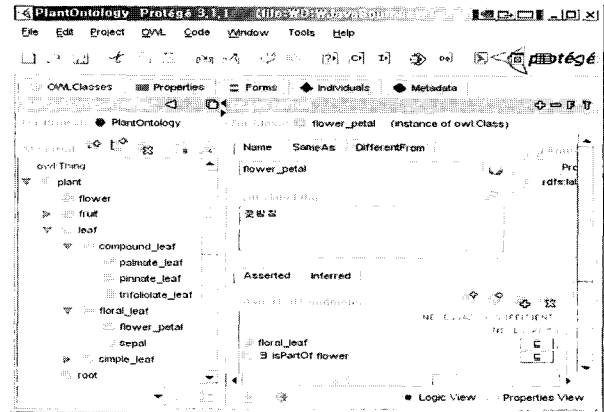


그림 13. Protégé를 이용하여 구축된 OWL ontology
Figure 13. OWL ontology developed by using Protégé

실험을 위해 과일(fruit)의 하위 개념을 편의상 10개로 제한하고, 각 개념 당 1개~2개의 학습 영상개체를 분석하여 개념의 속성 범위를 정하였다. 그리고 배경과 객체가 쉽게 구분되는 단일 영상을 대상으로 실험하였다. 그림 14와 같이 각 개념별로 2개의 영상을 학습 영상으로 선정하였고, 학습에 사용되지 않은 59개의 테스트 영상을 영상이해 실험 대상으로 하였다. 테스트 영상에 대한 영상이해 실험 결과를 표 2에 보인다.

실험에서 사용된 변수는 $d' = 10^\circ$, $r' = 10^\circ$, $L = 10$, $\alpha = 0.3$, $\beta = 0.3$, $\gamma = 0.3$ 과 같다.

표 2에서 1개의 영상개체를 이용하여 영상개체의 특징 속성을 정의하였을 때 보다는 2개의 영상개체를 이용하여 좀 더 넓은 범위의 특징 속성을 정의하였을 때 오히려 정확률이 낮아지는 것을 확인할 수 있었다.

표 2와 같은 결과를 분석하기 위해 그림 15와 같은 "apple"개념의 영상에 대한 영상이해 실험 결과를 분석한다.

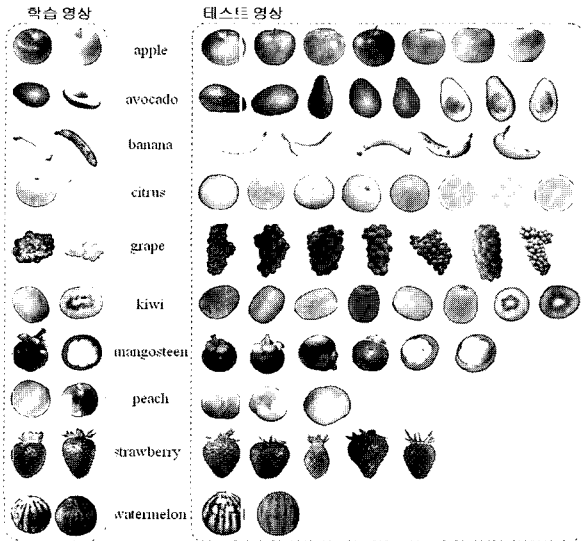


그림 14. 실험에 사용된 과일 영상
Figure 14. Images of fruits for tests

그림 15의 영상에 대한 영상이해 실험 결과인 표 3과 표 4는 각각 1개 영상과 2개의 영상을 학습했을 때의 각 테스트 영상에 대한 개념 후보를 나타낸 것이다. 그림 15에서 apple_ax(x=1,2,3,...)는 붉은색 계열의 사과이고 apple_bx는 녹색 계열의 사과이다.

붉은색 계열의 사과 영상인 apple_aL을 학습한 영상이해 실험의 결과인 표 3을 보면, apple_ax보다 apple_bx에서의 "apple" 개념의 순위가 상위에 위치함을 알 수 있다. 이것은 학습 영상인 apple_aL의 색상 속성이 붉은색 계열이고 테스트 영상인 apple_bx의 색상 속성은 녹색 계열이므로, apple_bx는 "kiwi", "avocado"와 같이 색상 속성이 녹색에 가까운 개념으로 해석될 수 있기 때문으로 판단된다.

붉은색 계열의 apple_aL과 녹색 계열의 apple_bL을 학습한 영상이해 실험의 결과인 표 4를 보면, apple_bx의 개념 후보에서 "apple" 개념이 상위에 위치한다는 것을 알 수 있다. 그러나 apple_a3의 경우에는 표 3의 결과보다 "apple" 개념의 순위가 낮게 나왔다. 이것은 그림 14에서 "peach" 개념

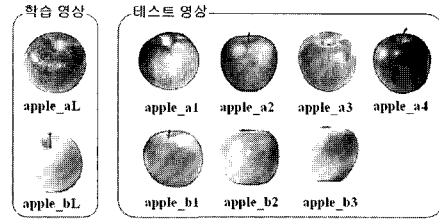


그림 15. 실험에 사용된 "apple" 영상
Figure 15. Images of "apple" for tests

표 2. 영상이해 실험 결과

Table 2. Results of tests of the image understanding

개념	1개 영상 학습 결과			2개 영상 학습 결과		
	실험영상개수	일치영상개수	정확률 (%)	실험영상개수	일치영상개수	정확률 (%)
apple	7	2	28.5	7	3	42.8
avocado	8	4	50	8	4	50
banana	5	1	20	5	1	20
citrus	8	8	100	8	8	100
grape	7	2	28.5	7	1	14.2
kiwi	8	6	75	8	3	37.5
mangosteen	6	3	50	6	3	50
peach	3	0	0	3	0	0
strawberry	5	5	100	5	4	80
watermelon	2	2	100	2	2	100
계	59	33	55.9	59	29	49.1

의 두 번째 학습 영상으로 사용된 영상의 속성들이 apple_a3 영상의 속성들과 흡사하여 발생한 경우이다. 그러나 apple_a3을 제외한 테스트 영상에서는 표 3보다 "apple" 개념의 순위가 전체적으로 상위에 위치함을 알 수 있다.

표 5와 그림 16은 각 테스트 영상의 해석 결과에서 유사도에 따라 후보 개념의 순위가 따른 누적 정확률(precision)을 보인다. 표 5에서 보는 바와 같이 영상 해석 후보 순위별 누적 정확률은 2개의 영상개체를 학습하였을 때 오히려 높아짐을 알 수 있다. 이것은 특정 속성의 범위가 넓게 정의될수록 영상개체를 해석할 수 있는 가능성이 높아지기 때문이라 판단된다.

표 3. apple_a2 학습에 대한 테스트 영상 후보 개념

Table 3. Concept candidates of test images by apple_a2 learning

apple_aL 학습

apple_a1 : [apple, peach, strawberry, grape, mangosteen, citrus, banana, watermelon, kiwi, avocado]

apple_a2 : [citrus, avocado, banana, peach, apple, mangosteen, kiwi, strawberry, grape, watermelon]

apple_a3 : [apple, strawberry, grape, banana, peach, watermelon, citrus, mangosteen, kiwi, avocado]

apple_a4 : [strawberry, apple, mangosteen, peach, grape, watermelon, banana, kiwi, citrus, avocado]

apple_b1 : [kiwi, avocado, citrus, banana, mangosteen, peach, apple, strawberry, grape, watermelon]

apple_b2 : [avocado, citrus, banana, apple, kiwi, peach, mangosteen, strawberry, grape, watermelon]

apple_b3 : [kiwi, avocado, banana, citrus, apple, mangosteen, peach, grape, strawberry, watermelon]

표 4. apple_a2와 apple_b4 학습에 대한 테스트 영상 개념 후보

Table 4. Concept candidates of test images by apple_a2 and apple_b4 learning

apple_aL 학습

apple_a1 : [apple, peach, strawberry, mangosteen, citrus, grape, banana, avocado, watermelon, kiwi]

apple_bL 학습

apple_a2 : [citrus, avocado, banana, apple, grape, peach, kiwi, mangosteen, strawberry, watermelon]

apple_a3 : [peach, apple, strawberry, mangosteen, grape, avocado, banana, citrus, watermelon, kiwi]

apple_a4 : [strawberry, apple, mangosteen, peach, grape, avocado, watermelon, banana, citrus, kiwi]

apple_b1 : [apple, banana, grape, citrus, avocado, kiwi, mangosteen, peach, watermelon, strawberry]

apple_b2 : [avocado, citrus, grape, apple, banana, kiwi, watermelon, mangosteen, peach, strawberry]

apple_b3 : [apple, banana, kiwi, grape, avocado, citrus, mangosteen, peach, watermelon, strawberry]

표 5. 영상 해석 순위별 누적 정확률 (%)
Table 5. The cumulative precision in ranking of image understanding (%)

순위	1	2	3	4	5	6	7	8	9
1개 영상개체 학습	55.9	67.7	74.5	81.3	89.8	91.5	96.6	98.3	100
2개 영상개체 학습	49.1	72.8	79.6	88.1	89.8	91.5	94.9	98.3	100

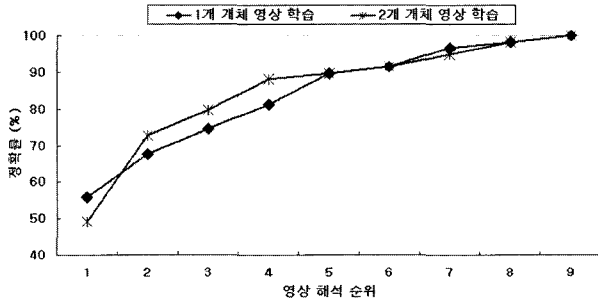


그림 16. 누적 정확률 그래프
Figure 16. A graph of cumulative precision

5. 결 론

본 논문에서는 온톨로지로 표현한 지식에 근거하여 영상의 처리, 분석, 해석 과정을 통해 영상을 이해하는 영상이해 프로세스와 시스템을 제안하였다. 그리고 제안된 영상 이해 프로세스에 기반 하여 영상이해 시스템을 구축하고, 특정 분야에서의 실험을 통하여 제안된 프로세스와 시스템의 효용성을 확인하였다.

영상 이해를 통해 영상이 표현하는 객체를 자동으로 인식함으로써 인간이 직접 처리할 수 없는 방대한 양의 영상 정보를 처리할 수 있게 된다. 따라서 영상 이해 시스템은 로봇 비전, 상황 인식, 영상 분류, 영상 검색 등과 같은 영상 관련 분야에 적극 활용될 수 있다.

향후 연구과제로서 첫째로는 본 논문에서 도입한 영상의 특징 속성, 즉, 객체의 색상, 모양, 형태 정보만으로는 객체를 충분히 설명할 수 없는 한계가 있으므로 이에 대한 보다 많은 연구가 이루어져야 할 것으로 판단된다. 둘째로는 영상 내 객체에 대한 해석뿐만 아니라 객체의 동작이나 상황을 알아내는 것이 영상이해의 목적중의 하나이므로 객체 사이의 관계 또한 객체 사이의 상호 작용을 정의한 온톨로지의 확대 적용에 관한 연구가 이루어져야 한다고 판단된다.

참 고 문 헌

[1] 손세호, 이인근, 권순학, "Ontological 지식 기반 영상이해시스템의 구조," KFIS Spring Conference, Vol. 14, No. 1, pp. 235-240, 2004.
[2] T. Matsuyama, V. Hwang, *SIGMA: A Knowledge-Based Aerial Image Understanding System*, Plenum, New York, 1990.
[3] D. Crevier and R. Lepage, "Knowledge-Based Image Understanding Systems: A Survey," *Computer Vision and Image Understanding*, Vol. 67, No. 2, pp 161-185, 1997.

[4] Jiebo Luo, Andreas E. Savakis, Amit Singhal, "A Bayesian network-based framework for semantic image understanding," *Pattern Recognition*, Vol. 38, Issue 6, pp. 919-934, 2005.
[5] N. Maillot, M. Thonnat, A. Boucher, "Toward Ontology Based Cognitive Vision," In *Proceedings of The Third International Conference On Computer Vision Systems*, LNCS 2626, pp. 44-53, 2003.
[6] N. Maillot, M. Thonnat, A. Boucher, "A Visual Concept Ontology for Automatic Image Recognition," In *Proceedings of The Second International Semantic Web Conference*, 2003.
[7] Natalya F. Noy and Deborah L. McGuinness, "Ontology Development 101: A Guide to Creating Your First Ontology," *Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880*, March 2001.
[8] OWL Web Ontology Language Guide, <http://www.w3.org/TR/owl-guide>, July 2006.
[9] Wordnet, <http://wordnet.princeton.edu>, October 2006.
[10] Protégé, <http://protege.stanford.edu>, October 2006.

저 자 소 개

이인근(In K. Lee)

제16권 6호(2006년 12월호) 참조

서석태(Suk Tae Seo)

제16권 6호(2006년 12월호) 참조

정혜천(Hye C. Jung)

제16권 6호(2006년 12월호) 참조



손세호(Seo. H. Son)

2000년 : 영남대학교 전기전자공학부 (공학사)

2002년 : 영남대학교 대학원 전기공학과 (공학석사)

2004년 : 영남대학교 대학원 전기공학과 박사과정수료

2004년~현재 : 탑 엔지니어링(주) 근무

관심분야 : 지능시스템 및 제어, 영상이해, 비전시스템
E-mail : kazma@lycos.co.kr

권순학(Soon H. Kwon)

제16권 6호(2006년 12월호) 참조